Hamilton-Jacobi-Bellman equations for Quantum Filtering and Control

J. Gough[a], V.P. Belavkin[b], and O.G. Smolyanov[c]
[a]Nottingham-Trent University, [b]Nottingham University,
[c]Moscow State Univeristy

Abstract

We exploit the separation of the filtering and control aspects of quantum feedback control to consider the optimal control as a classical stochastic problem on the space of quantum states. We derive the corresponding Hamilton-Jacobi-Bellman equations using the elementary arguments of classical control theory and show that this is equivalent, in the Stratonovich calculus, to a stochastic Hamilton-Pontryagin setup. We show that, for cost functionals that are linear in the state, the theory yields the traditional Bellman equations treated so far in quantum feedback. A controlled qubit with a feedback is considered as example.

1 Introduction

When engineers set about to control a classical system, they can evoke the celebrated *Separation Theorem* which allows them to treat the problem of estimating the state of the system (based on typically partial observations) from the problem of how to optimally control the system (through feedback of these observations into the system dynamics), see for instance [13]. Remarkably, as it was pointed out for the first time in [2], this is also true when trying to control the quantum world, see also [3],[8],[11]. To begin with, the very act of measurement itself never supplies anything but incomplete information about the state of a system and, as is well known, alters the state in process. However, provided we use a non-demolition principle [3] when measuring the system over time, we can apply a filter scheme for state estimation continuously in time. The general theory of the continuous in time nondemolition measurements and filtering was developed by Belavkin in [3],[5],[6],[7], however we will use here its final result for a simple quantum diffusion model described by the quantum state filtering equation with a single white noise innovation, see e.g. [4],[29],[12]. We should emphasize that the continuous-time filtering theory for this case can be obtained as the limit of a discrete-time measurements where nothing beyond the standard von Neumann projection postulate is used [19],[20], [25], [26]. Once the filtered dynamics is known, the of optimal feedback control of the system can then be

formulated as a distinct problem. Modern experimental physics has opened up unprecedented opportunities to manipulate the quantum world, and feedback control has been already been successfully implemented for real physical systems [1],[18]. Currently, these activities have attracted interest in the related mathematical issues such as stability, observability, etc., [11],[21],[15],[22].

The separation of the classical world from the quantum world is, in practice, the most notoriously troublesome task faced in modern physics. At the very heart of this issue is the very different meanings we attach to the word *state*. What we want to remark upon, and exploit, is the fact that the separation of the control problem from the filtering gives us just the required separation of classical from quantum features. By the quantum state we mean the von Neumann density matrix which yields all the (stochastic) information available about the system at the current time - this we also take to be the state in the sense used in control engineering. All the quantum features are contained in this state, and the filtering equation it satisfies may then to be understood as classical stochastic differential equation which just happens to have solutions that are von Neumann density matrix valued stochastic processes. The ensuing problem of determining optimal control may then be viewed as a classical problem, albeit on the unfamiliar state space of von Neumann density matrices rather than the Euclidean spaces to which we are usually accustomed. Once we get used to this setting, the problem of dynamical programming, Bellman's optimality principle, and so on, can be formulated in the same spirit as before.

We shall consider optimization for cost functions that are non-linear functionals of the state. Traditionally quantum control has been restricted to linear functions where - given the physical meaning attached to a quantum state - the cost functions are therefore expectations of certain observables. In this situation, which we consider as a special case, we see that the distinction between classical and quantum features may be blurred: that is, the classical information about the measurement observations can be incorporated as additional randomness into the quantum state. This is the likely reason why the separation does not seem to have been taken up before.

2 Notations

The Hilbert space for our fixed quantum system will be a complex, separable Hilbert space $\mathfrak{h}$. We shall use the following spaces of operators:

$\mathcal{A} = \mathfrak{B}(\mathfrak{h})$	- the Banach algebra of bounded operators on $\mathfrak{h}$;
$\mathcal{A}_\star = \mathfrak{I}(\mathfrak{h})$	- the predual space of trace-class operators on $\mathfrak{h}$;
$\mathcal{S} = \mathfrak{S}(\mathfrak{h})$	- the positive, unital trace operators (states) on $\mathfrak{h}$.

The space $\mathcal{A}_\star$ equipped with the trace norm $\|\varrho\|_1 = \mathrm{tr}\,|\varrho|$ is the complex Banach space, the dual of which is identified with the algebra $\mathcal{A}$ with usual operator norm. The natural duality between the spaces $\mathcal{A}_\star$ and $\mathcal{A}$ is indicated by

$$\langle \varrho, X \rangle := \mathrm{tr}\,\{\varrho X\}\,, \tag{1}$$

for each $\varrho \in \mathcal{A}_\star, X \in \mathcal{A}$. The positive elements of $\mathcal{A}_\star$ normalized as $\|\varrho\|_1 = 1$ are called normal states, and the extremal elements $\varrho \in \mathcal{S}$ of the convex set $\mathcal{S} \subset \mathcal{A}_\star$ correspond to pure quantum states. The symmetric tensor power $\mathcal{A}_{sym}^{\otimes 2} = \mathcal{A} \otimes_{sym} \mathcal{A}$ of the algebra $\mathcal{A}$ is the subalgebra of $\mathfrak{B}\left(\mathfrak{h}^{\otimes 2}\right)$ of all bounded operators on the Hilbert product space $\mathfrak{h}^{\otimes 2} = \mathfrak{h} \otimes \mathfrak{h}$, commuting with the unitary involutive operator $S = S^\dagger$ of permutations $\eta_1 \otimes \eta_2 \mapsto \eta_2 \otimes \eta_1$ for any $\eta_i \in \mathfrak{h}$.

A map $\mathcal{L}(\cdot)$ from $\mathcal{A} = \mathfrak{B}(\mathfrak{h})$ to itself is said to be a Lindblad generator if it takes the form

$$\mathcal{L}(X) = -i[X, H] + \sum_\alpha \mathcal{L}_{R_\alpha}(X), \tag{2}$$

$$\mathcal{L}_R(X) = R^\dagger X R - \frac{1}{2} R^\dagger R X - \frac{1}{2} X R^\dagger R \tag{3}$$

with H self-adjoint, the $R_\alpha \in \mathcal{A}$ and $\sum_\alpha R_\alpha^\dagger R_\alpha$ (ultraweakly convergent [23] for an infinite set $\{R_\alpha\}$). The generator is Hamiltonian if it just takes form $i[H, \cdot]$. The preadjoint $\mathcal{L}' = \mathcal{L}_\star$ of a generator $\mathcal{L}$ is defined on the preadjoint space $\mathcal{A}_\star$ through the relation $\langle \mathcal{L}'(\varrho), X \rangle = \langle \varrho, \mathcal{L}(X) \rangle$. We note that Lindblad generators have the property $\mathcal{L}(I) = 0$ corresponding to conservation of the identity operator $I \in \mathcal{A}$ or, equivalently, $\mathrm{tr}\{\mathcal{L}'(\varrho)\} = 0$ for all $\varrho \in \mathcal{A}_\star$.

In quantum control theory it is necessary to consider time-dependent generators $\mathcal{L}(t)$, through an integrable time dependence of the controlled Hamiltonian $H(t)$, or more generally due to a square-integrable time dependence of the coupling operators $R_\alpha(t)$. We will always assume that these integrability conditions, corresponding to the existence of the unique solution $\varrho(t) = P_t(t_0, \varrho_0)$ to the quantum state Master equation

$$\frac{d}{dt}\varrho(t) = \mathcal{L}'(t, \varrho(t)) \equiv v(t, \varrho(t)), \tag{4}$$

for all for $t \geq t_0$, given an initial condition $\varrho(t_0) = \varrho_0 \in \mathcal{S}$, are fulfilled.

Let $\mathsf{F} = \mathsf{F}[\cdot]$ be a (nonlinear) functional $\varrho \mapsto \mathsf{F}[\varrho]$ on $\mathcal{S}$, then we say it admits a (Fréchet) derivative if there exists a $\mathcal{A}$-valued function $\nabla_\varrho \mathsf{F}[\cdot]$ on $\mathcal{A}_\star$ such that

$$\lim_{h \to 0} \frac{1}{h} \{\mathsf{F}[\cdot + h\tau] - \mathsf{F}[\cdot]\} = \langle \tau, \nabla_\varrho \mathsf{F}[\cdot] \rangle, \tag{5}$$

for each $\tau \in \mathcal{A}_\star$. In the same spirit, a Hessian $\nabla_\varrho^{\otimes 2} \equiv \nabla_\varrho \otimes \nabla_\varrho$ can be defined as a mapping from the functionals on $\mathcal{S}$ to the $\mathcal{A}_{sym}^{\otimes 2} := \mathcal{A} \otimes_{sym} \mathcal{A}$-valued functionals, via

$$\lim_{h, h' \to 0} \frac{1}{hh'} \{\mathsf{F}[\cdot + h\tau + h'\tau'] - \mathsf{F}[\cdot + h\tau] - \mathsf{F}[\cdot + h'\tau'] + \mathsf{F}[\cdot]\} = \langle \tau \otimes \tau', \nabla_\varrho \otimes \nabla_\varrho \mathsf{F}[\cdot] \rangle. \tag{6}$$

and we say that the functional is twice continuously differentiable whenever $\nabla_\varrho^{\otimes 2} \mathsf{F}[\cdot]$ exists and is continuous in the trace norm topology.

Likewise, a functional $f : X \mapsto f[X]$ on $\mathcal{A}$ is said to admit an $\mathcal{A}_\star$-derivative if there exists an $\mathcal{A}_\star$-valued function $\nabla_X f[\cdot]$ on $\mathcal{A}$ such that

$$\lim_{h\to 0} \frac{1}{h} \{f[\cdot + hA] - f[\cdot]\} = \langle \nabla_X f[\cdot], A\rangle \tag{7}$$

for each $A \in \mathfrak{B}(\mathfrak{h})$.

With the customary abuses of differential notation, we have for instance

$$\nabla_\varrho f(\langle \varrho, X\rangle) = f'(\langle \varrho, X\rangle) X, \quad \nabla_X f(\langle \varrho, X\rangle) = f'(\langle \varrho, X\rangle) \varrho.$$

Typically, we shall use ∇_ϱ more often, and tend denote it by δ (as "inverse" to the notation ϱ), leaving the simple notation ∇ for ∇_X.

3 Quantum Filtering Equation

The state of an individual continuously measured quantum system does not coincide with the solution $\varrho(t)$ of the deterministic master equation (4) but is a $\mathcal{S}$-valued stochastic process $\varrho_\bullet(t) : \omega \mapsto \varrho_\omega(t)$ which depends on the random measurement output $\omega = \{\omega(t)\} \in \Omega$ in a causal manner. We take the output process to constitute a white noise, in which case we may work with the innovations process which will be a Wiener process $W(t)$ defined in the generalized sense by $\frac{d}{dt}W(t) = \omega(t)$ with $W(0) = 0$. The Belavkin quantum filtering equation in this case is [4],[7], [29],[12]

$$d\varrho_\bullet(t) = w(t, u(t), \varrho_\bullet(t))\, dt + \sigma(\varrho_\bullet(t))\, dW(t) \tag{8}$$

where $dW(t) = W(t+dt) - W(t)$, the time coefficient is

$$w(t, u, \varrho) = i[\varrho, H(t,u)] + \mathcal{L}'_R(\varrho) + \mathcal{L}'_L(\varrho), \tag{9}$$

with $\mathcal{L}'_L(\varrho)$ of the form given

$$\mathcal{L}'_L(\varrho) = L\varrho L^\dagger - \frac{1}{2}\varrho L^\dagger L - \frac{1}{2}L^\dagger L\varrho,$$

and the fluctuation coefficient is

$$\sigma(\varrho) = L\varrho + \varrho L^\dagger - \langle \varrho, L + L^\dagger\rangle \varrho. \tag{10}$$

Here L is a bounded operator describing the coupling of the system to the measurement apparatus.

The time coefficient w consists of three separate terms: The first term is Hamiltonian and depends on a control parameter u belonging to some parameter space $\mathcal{U}$ which we must specify at each time; the second term is the adjoint of a general Lindblad generator $\mathcal{L}_R$ and describes the uncontrolled, typically dissipative, effect of the environment; the final term is adjoint to the Lindblad generator $\mathcal{L}_L(X)$ which is related to the coupling operator L.

The maps w and σ are required to be Lipschitz continuous in all their components: for L constant and bounded, this will be automatic for the ϱ-variable with the notion of trace norm topology. We remark that $\mathrm{tr}\left\{\sigma\left(\varrho\right)\right\}=0$ and, by conservativity, $\mathrm{tr}\left\{w\left(t,u,\varrho\right)\right\}=0$ for all $\varrho\in\mathcal{A}_{\star}$. This implies that the normalization $\mathrm{tr}\left\{\varrho\right\}$ is a conserved quantity $\mathrm{tr}\left\{\varrho_{\bullet}\left(t\right)\right\}=\mathrm{tr}\left\{\varrho\right\}$ under the stochastic evolution (8).

A choice of control function $\left\{u\left(t\right):t\in\left[T_1,T_2\right]\right\}$ is required before we can solve the filtering equation (8) on the time interval $\left[T_1,T_2\right]$ for given initial state at time T_1. From what we have said above, this is required to be a $\mathcal{U}$-valued function which we take to be continuous for the moment.

Let $\left\{P_{r,\omega}\left(t,\varrho\right):r\geq t,\omega\in\Omega\right\}$ be the solution $\varrho_{\bullet}\left(r\right)=P_{r,\bullet}\left(t,\varrho\right)$ to (8) starting in state $\varrho_{\omega}\left(t\right)=\varrho$ at time $r=t$ for all $\omega\in\Omega$. This will be a Markov process in $\mathcal{S}$ (embedded in the Banach space $\mathcal{A}_{\star}$), see for instance [14], and we remark that, for twice continuously differentiable functionals F on $\mathcal{A}_{\star}$, we will have

$$\lim_{h\rightarrow0^{+}}\frac{1}{h}\left\{\mathbb{E}\left[\mathsf{F}\left[P_{t+h,\bullet}\left(t,\varrho\right)\right]-\mathsf{F}\left[\varrho\right]\right]\right\}=D\left(t,u,\varrho\right)\mathsf{F}\left[\varrho\right],$$

where $D\left(t,u,\varrho\right)$ is the elliptic operator defined by

$$D\left(t,u,\varrho\right)\cdot=\left\langle w\left(t,u,\varrho\right),\delta\cdot\right\rangle+\frac{1}{2}\left\langle\sigma\left(\varrho\right)\otimes\sigma\left(\varrho\right),\left(\delta\otimes\delta\right)\cdot\right\rangle. \tag{11}$$

For the classical analogue of stochastic flows on manifolds, see for instance [10].

3.1 Stratonovich Version

We convert to the Stratonovich picture [27] by means of the identity [16]

$$\sigma\left(\varrho_{\bullet}\right)\,dW=\sigma\left(\varrho_{\bullet}\right)\circ dW-\frac{1}{2}d\sigma\left(\varrho_{\bullet}\right).dW$$

and from (10) we get

$$d\sigma\left(\varrho_{\bullet}\right)=Ld\varrho_{\bullet}+d\varrho_{\bullet}L^{\dagger}-\left\langle d\varrho_{\bullet},L+L^{\dagger}\right\rangle\varrho_{\bullet}-\left\langle\varrho_{\bullet},L+L^{\dagger}\right\rangle d\varrho_{\bullet}-\left\langle d\varrho_{\bullet},L+L^{\dagger}\right\rangle d\varrho_{\bullet}.$$

After a little algebra, we obtain the Stratonovich form of the Belavkin filtering equation:

$$d\varrho_{\bullet}=v\left(t,u,\varrho_{\bullet}\right)\,dt+\sigma\left(\varrho_{\bullet}\right)\circ dW \tag{12}$$

where, with $\sigma\equiv\sigma\left(\varrho\right)$,

$$\begin{aligned}v\left(t,u,\varrho\right)&=w\left(t,u,\varrho\right)-\frac{1}{2}\left\{L\sigma+\sigma L^{\dagger}-\left\langle\sigma,L+L^{\dagger}\right\rangle\varrho-\left\langle\varrho,L+L^{\dagger}\right\rangle\sigma\right\}\\&=i\left[\varrho,H\left(t,u\right)\right]+\mathcal{L}_{R}'\left(\varrho\right)+\Big\{K\left(\varrho\right)\,\varrho+\varrho\,K\left(\varrho\right)^{\dagger}+\mathsf{F}\left(\varrho\right)\,\varrho\Big\}\end{aligned} \tag{13}$$

where we introduce the operator-valued function

$$K\left(\varrho\right):=-\frac{1}{2}\left(L+L^{\dagger}\right)L+\left\langle\varrho,L+L^{\dagger}\right\rangle L \tag{14}$$

and the scalar-valued function

$$\mathsf{F}(\varrho) := \frac{1}{2}\left\langle \varrho, L^2 + 2L^\dagger L + L^{\dagger 2}\right\rangle - \left\langle \varrho, L + L^\dagger \right\rangle^2 . \tag{15}$$

We refer to w in (9) and v in (13) as the *Itô and Stratonovich state velocities*, respectively. We note that the decoherent component $L\varrho L^\dagger$ appearing in $\mathcal{L}_L'$, and present in $w(t,u,\varrho)$, is now absent in $v(t,u,\varrho)$.

The elliptical operator $D(t,u,\varrho)$ can then be put into Hörmander form as

$$D(t,u,\varrho)(\cdot) := \langle v(t,u,\varrho), \delta\cdot\rangle + \frac{1}{2}\langle \sigma(\varrho), \delta\langle \sigma(\varrho), \delta\cdot\rangle\rangle , \tag{16}$$

by using the equality (13) in the definition (11).

4 Optimal Control

The cost for a control function $\{u(r)\}$ over any time-interval $[t,T]$ is random, taken to have the integral form

$$\mathsf{J}_\omega[\{u(r)\};t,\varrho] = \int_t^T \mathsf{C}(r,u(r),\varrho_\omega(r))\,dr + \mathsf{S}(\varrho_\omega(T)) \tag{17}$$

where $\{\varrho_\bullet(r) : r\in[t,T]\}$ is the solution to the filtering equation with initial condition $\varrho_\bullet(t) = \varrho$. We assume that the cost density C and the terminal cost S will be continuously differentiable in each of its arguments. In fact, due to the statistical interpretation of quantum states, we should consider only the linear dependence

$$\mathsf{C}(r,u,\varrho) = \langle \varrho, C(r,u)\rangle,\ \mathsf{S}(\varrho) = \langle \varrho, S\rangle \tag{18}$$

of C and S on the state ϱ as it was already suggested in [2],[3],[8]. We will explicitly consider this case later, but for the moment we will not use the linearity of C and S.

The feedback control $u(t)$ is to be considered a random variable $u_\omega(t)$ adapted with respect to the innovation process $W(t)$ and so we therefore consider the problem of minimizing its average cost value with respect to $\{u_\bullet(t)\}$. To this end, we define the optimal average cost to be

$$\mathsf{S}(t,\varrho) := \inf_{\{u_\bullet(r)\}} \mathbb{E}\left[\mathsf{J}_\bullet[\{u_\bullet(r)\};t,\varrho]\right], \tag{19}$$

where the minimum is considered over all measurable adapted control strategies $\{u_\bullet(r) : r\geq t\}$. The aim of feedback control theory is then to find an optimal control strategy $\{u_\bullet^*(t)\}$ and evaluate $\mathsf{S}(t,\varrho)$ on a fixed time interval $[t_0,T]$. Obviously that the cost $\mathsf{S}(t,\varrho)$ of the optimal feedback control is in general smaller then the minimum of $\mathbb{E}[\mathsf{J}_\bullet[\{u\};t,\varrho]]$ over nonstochastic strategies $\{u(r)\}$ only, which gives the solution of the open loop (without feedback) quantum control problem. In the case of the linear costs (18) this open-loop problem is equivalent to the following quantum deterministic optimization problem which can be tackled by the classical theory of optimal deterministic control in the corresponding Banach spaces.

4.1 Bellman & Hamilton-Pontryagin Optimality

Let us first consider nonstochastic quantum optimal control theory assuming that the state $\varrho(t) \in \mathcal{S}$ obeys the master equation (4) where $v(t,\cdot) = \mathcal{L}'(t, u(t), \cdot)$ is an adjoint of some Lindblad generator $\mathcal{L}'(t, u, \cdot) \equiv v(t, u, \cdot)$ for each t and u with, say, the control being exercised in the Hamiltonian component $i[\cdot, H(t, u)]$ as before. The control strategy $\{u(t)\}$ will be here non-random, as will be any specific cost $J[\{u\}; t_0, \varrho_0]$. For times $t < t+\varepsilon < T$, one has

$$\mathsf{S}(t,\varrho) = \inf_{\{u\}} \left\{ \int_t^{t+\varepsilon} \mathsf{C}(r, \varrho(r), u(r))\, dr + \int_{t+\varepsilon}^{T} \mathsf{C}(r, \varrho(r), u(r))\, dr + \mathsf{S}(\varrho(T)) \right\}.$$

Suppose that $\{u^*(r) : r \in [t, T]\}$ is an optimal control when starting in state ϱ at time t, and denote by $\{P_r(t,\varrho) : r \in [t,T]\}$ the corresponding state dynamics $\varrho^*(r) = P_r(t,\varrho)$, $P_t = \varrho$. Bellman's optimality principle [9],[13] observes that the control $\{u^*(r) : r \in [t+\varepsilon, T]\}$ will then be optimal when starting from $\varrho^*(t+\varepsilon)$ at the later time $t+\varepsilon$. It therefore follows that

$$\mathsf{S}(t,\varrho) = \inf_{\{u(r)\}} \left\{ \int_t^{t+\varepsilon} \mathsf{C}(r, u(r), \varrho(r))\, dr + \mathsf{S}(t+\varepsilon, \varrho(t+\varepsilon)) \right\}.$$

For ε small we expect that $\varrho(t+\varepsilon) = \varrho + v(t, u(t), \varrho)\varepsilon + o(\varepsilon)$ and provided that S is sufficiently smooth we may make the Taylor expansion

$$\mathsf{S}(t+\varepsilon, \varrho(t+\varepsilon)) = \left[1 + \varepsilon \frac{\partial}{\partial t} + \varepsilon \langle v(t, u(t), \varrho), \delta \rangle \right] \mathsf{S}(t,\varrho) + o(\varepsilon). \tag{20}$$

In addition, we approximate

$$\int_t^{t+\varepsilon} \mathsf{C}(r, u(r), \varrho(r))\, dr = \varepsilon \mathsf{C}(t, u(t), \varrho) + o(\varepsilon)$$

and conclude that

$$\mathsf{S}(t,\varrho) = \inf_{u \in U} \left\{ \left[1 + \varepsilon \left(\mathsf{C}(t,u,\varrho) + \frac{\partial}{\partial t} + \langle v(t,u,\varrho), \delta \rangle \right) \right] \mathsf{S}(t,\varrho) \right\} + o(\varepsilon)$$

where now the infimum is taken over the point-value of $u(t) = u \in U$. In the limit $\varepsilon \to 0$, one obtains the equation

$$\frac{\partial}{\partial t} \mathsf{S}(t,\varrho) + \inf_{u \in \mathcal{U}} \{ \mathsf{C}(t,u,\varrho) + \langle v(t,u,\varrho), \nabla \mathsf{S}(t,\varrho) \rangle \} = 0. \tag{21}$$

The equation is then to be solved subject to the terminal condition

$$\mathsf{S}(T,\varrho) = \mathsf{S}(\varrho). \tag{22}$$

We may introduce the Pontryagin Hamiltonian function on $[0,T] \times \mathcal{S} \times \mathcal{A}$ defined by the Legendre-Frenchel transform

$$\mathcal{H}_v(t,\varrho,X) := \sup_{u \in \mathcal{U}} \{ \langle v(t,u,\varrho), \lambda I - X \rangle - \mathsf{C}(t,u,\varrho) \}, \tag{23}$$

(which in fact does not depend on $\lambda \in \mathbb{C}$ since $\langle v(t,u,\varrho), I\rangle = 0$). It should be emphasized that these Hamiltonians are purely classical devices which may be called super-Hamiltonians to be distinguished from H. We may then rewrite (21) as the (backward) *Hamilton-Jacobi* equation

$$\frac{\partial}{\partial t}\mathsf{S}(t,\varrho) = \mathcal{H}_v(t,\varrho,\delta\mathsf{S}(t,\varrho)). \tag{24}$$

The operator-valued function $X(t,\varrho) = \delta\mathsf{S}(t,\varrho)$ satisfying then the equation $\frac{d}{dt}X = \delta\mathcal{H}_v(t,\rho,X)$ is referred to as the *co-state*, with the terminal condition $X(T,\varrho) = \delta\mathsf{S}(\varrho)$. We remark that, if $u^*(t,\varrho,X)$ is an optimal control minimizing

$$\mathcal{K}_v(t,u,\varrho,X) = \langle v(t,u,\varrho), \lambda I - X\rangle - \mathsf{C}(t,u,\varrho),$$

then the corresponding state dynamical equation $\frac{d}{dt}\varrho = v(t,u^*(t,\varrho,X),\varrho)$ in terms of its optimal solution $P_t \equiv P_t(t_0,\varrho_0)$ corresponding to $P_{t_0} = \varrho^*(t_0) \equiv \varrho_0$ can be written as $\dot{P} = -\nabla_Q\mathcal{H}_v(t,P,Q)$ noting that

$$\mathcal{H}_v(t,P,Q) = \langle v(t,u^*(t,P,Q),P), \lambda I - Q\rangle - \mathsf{C}(t,u^*(t,P,Q),P),$$

where $Q_t = X(t)$ is the solution $X(t) = Q_t(T,S)$ of $\dot{Q} = \nabla_P\mathcal{H}_v(t,P,Q)$ corresponding to $Q_T = \delta\mathsf{S}(\varrho) \equiv S$. Thus we may equivalently consider the system of Hamiltonian equations

$$\begin{cases} \dot{P}_t + \nabla_Q\mathcal{H}_v(t,P_t,Q_t) = 0, \\ \dot{Q}_t - \nabla_P\mathcal{H}_v(t,P_t,Q_t) = 0. \end{cases} \tag{25}$$

which we refer to as the *Hamilton-Pontryagin equations*, in direct analogy with the classical case [24]. If we set $u^* = u^*(t,P,Q)$ such that $\mathcal{K}_v(t,u^*,P,Q) = \sup_{u\in\mathcal{U}}\mathcal{K}_v(t,u,P,Q)$, then the Pontryagin maximum principle is the observation that, for state and co-state $\{P\}$ and $\{Q\}$ respectively leading to optimality, we will have $\mathcal{K}_v(t,u,P,Q) \leq \mathcal{H}_v(t,P,Q)$ with equality for $u = u^*(t,P,Q)$ maximizing $\mathcal{K}_v(t,u,P,Q)$.

4.2 Bellman Equation for Filtered Dynamics

We now consider the stochastic differential equation (8) for the filtered state in place of the master equation (4). This time, the cost is random and we consider the problem of computing the minimum average cost as in (19). The Bellman principle can however be applied once more. As before, we let $\{u^*_\omega(t)\}$ be a stochastic adapted control leading to optimality and let $\varrho^*_\omega(r) = P_{r,\omega}(t,\varrho)$ be the corresponding state trajectory (now a stochastic process) starting from ϱ at time t. Again choosing $t < t+\varepsilon < T$, we have by the Bellman principle

$$\mathbb{E}\left[\mathsf{S}(t+\varepsilon,\varrho^*_\bullet(t+\varepsilon))\right] = \mathsf{S}(t,\varrho)$$
$$+\inf_{u\in\mathcal{U}}\mathbb{E}\left\{\frac{\partial\mathsf{S}}{\partial t}(t,\varrho) + \mathsf{C}(t,u,\varrho) + D(t,u,\varrho)\,\mathsf{S}(t,\varrho)\right\}\varepsilon + o(\varepsilon)$$

Taking the limit $\varepsilon \to 0$ yields the diffusive Bellman equation

$$\frac{\partial \mathsf{S}}{\partial t} + \inf_{u \in \mathcal{U}} \left\{ \mathsf{C}\left(t, u, \varrho\right) + D\left(t, u, \varrho\right) \mathsf{S}\left(t, \varrho\right) \right\} = 0.$$

This equation to be solved backward with the terminal condition $\mathsf{S}\left(T, \varrho\right) = \mathsf{S}\left(\varrho\right)$. Using the super-Hamiltonian function

$$\mathcal{H}_w\left(t, \varrho, X\right) := \sup_{u \in \mathcal{U}} \left\{ \left\langle w\left(t, u, \varrho\right), \lambda I - X \right\rangle - \mathsf{C}\left(t, u, \varrho\right) \right\}$$

this can be written either in the Hamilton-Jacobi form as

$$\frac{\partial \mathsf{S}}{\partial t} + \frac{1}{2} \left\langle \sigma\left(\varrho\right) \otimes \sigma\left(\varrho\right), \left(\delta \otimes \delta\right) \mathsf{S} \right\rangle = \mathcal{H}_w\left(t, \varrho, \delta \mathsf{S}\right). \tag{26}$$

5 Stochastic Hamilton-Jacobi-Bellman Equation

An alternative approach to deriving the equation (26) will now be formulated. First of all we make a Wong-Zakai approximation [30] to the Stratonovich filtering equation (12). This is achieved by introducing a differentiable process $W_\omega^{(\lambda)}\left(t\right) = \int_0^t \omega^{(\lambda)}\left(r\right) dr$ converging to the Wiener noise $W_\omega\left(t\right)$ as $\lambda \to 0$ almost surely and uniformly for $t \in [0, T]$. We may then expect the same type of convergence for $\left\{ \varrho_\omega^{(\lambda)}\left(t\right) \right\}$, the solution to the random ODE

$$\frac{d}{dt} \varrho_\omega^{(\lambda)}\left(t\right) = v\left(t, u\left(t\right), \varrho_\omega^{(\lambda)}\left(t\right)\right) + \sigma\left(\varrho_\omega^{(\lambda)}\left(t\right)\right) \omega^{(\lambda)}\left(t\right)$$

with non random initial condition $\varrho_\omega^{(\lambda)}\left(t_0\right) = \varrho_0$, to the solution $\left\{ \varrho_\omega\left(t\right) \right\}$ with the same initial data $\varrho_\omega\left(t_0\right) = \varrho_0$.

If we fix the output $\omega \in \Omega$, then we have an equivalent non-random dynamical system for which we will have a minimal cost function and we denote this as $\mathsf{S}_\omega^{(\lambda)}\left(t_0, \varrho_0\right)$. Note that this depends on the assumed realization of the measurement output process and on the approximation parameter λ. The HJB equation for $\mathsf{S}_\omega^{(\lambda)}\left(t, \varrho\right)$ will be (24) with $v\left(t\right)$ now replaced by $v\left(t\right) + \sigma \omega^{(\lambda)}\left(t\right)$:

$$\frac{\partial}{\partial t} \mathsf{S}_\omega^{(\lambda)} + \left\langle \sigma\left(\varrho\right), \delta \mathsf{S}_\omega^{(\lambda)} \right\rangle \omega^{(\lambda)}\left(t\right) = \mathcal{H}_v\left(t, \varrho, \delta \mathsf{S}_\omega^{(\lambda)}\right)$$

Since $\sigma\left(\varrho\right) \omega^{(\lambda)}\left(t\right)$ doesn't depend on u, the corresponding optimal strategy $u_\omega^*\left(t\right)$ as the solution of the optimization problem

$$\inf_{u \in \mathcal{U}} \left\{ \mathsf{C}\left(u, \varrho\right) + \left\langle v\left(u, \varrho\right) + \sigma\left(\varrho\right) \omega^{(\lambda)}, X \right\rangle \right\} = \left\langle \sigma\left(\varrho\right) \omega^{(\lambda)}, X \right\rangle - \mathcal{H}_v\left(\varrho, X\right)$$

is the same function $u^*\left(t, \varrho, X\right)$ of $\varrho = \varrho_\omega^{(\lambda)}\left(t\right)$ and $X = \delta \mathsf{S}_\omega^{(\lambda)}$, independent of $\omega^{(\lambda)}\left(t\right)$. In the limit $\lambda \to 0$ we obtain the Stratonovich SDE

$$d\mathsf{S}_\omega\left(t, \varrho\right) + \left\langle \sigma\left(\varrho\right), \delta \mathsf{S}_\omega\left(t, \varrho\right) \right\rangle \circ dW\left(t\right) = \mathcal{H}_v\left(t, \varrho, \delta \mathsf{S}_\omega\left(t, \varrho\right)\right) dt. \tag{27}$$

which may be called a stochastic Hamilton-Jacobi-Bellman equation.

5.1 Interpretation of the Stochastic HJB equation

The expression $\mathsf{S}_\omega(t_0, \varrho_0)$ gives the optimal cost from start time t_0 to terminal time T when we begin in state ϱ_0 and have measurement output $\omega \in \Omega$. It evidently depends on the information $\{\omega(r) : r \in [t_0, T]\}$ only and is statistically independent of the noise $W(t) = \omega_t$ prior to time t_0. In this sense, the stochastic action $\mathsf{S}_\omega(t, \varrho)$ is **backward-adapted**. This point is of crucial importance: it means that the stochastic Hamiltonian-Jacobi-Bellman theory is not related directly to the stochastic Hamilton-Jacobi theory [28] where both the state and the action are always taken as be forward-adapted; it also means that we need to be careful when converting (27) to Itô form. This is a direct consequence of the fact that Bellman's principle works by backward induction.

Let us introduce the following time-reversed notations

$$\tau := T - t,\ \tilde{W}(\tau) := W(T-\tau) = W(t) \text{ and } \tilde{\mathsf{S}}_\omega(\tau, \varrho) := \mathsf{S}_\omega(T-\tau, \varrho) = \mathsf{S}_\omega(t, \varrho).$$

The process $\tau \mapsto \tilde{\mathsf{S}}_\bullet(\tau, \varrho)$ is forward adapted to the filtration generated by $\tilde{W}$: that is $\tilde{\mathsf{S}}_\bullet(\tau, \varrho)$ is measurable with respect to the sigma algebra generated by $\{\tilde{W}(\sigma)\colon \sigma \in [0, \tau]\}$. Note that the Itô differential $d\tilde{W}(\tau) = \tilde{W}(\tau + \varepsilon) - \tilde{W}(\tau)$ coincides with $W(t-\varepsilon) - W(t) \equiv -\tilde{d}W(t)$ for $t = T - \tau \equiv \tilde{\tau}$.

Theorem 1 *The stochastic process $\{\mathsf{S}_\bullet(t, \varrho) : t \in [0, T]\}$ satisfies the backward Itô SDE*

$$d\mathsf{S}_\bullet + \frac{1}{2}\langle \sigma, \nabla \langle \sigma, \nabla \mathsf{S}_\bullet \rangle\rangle dt + \langle \sigma, \nabla \mathsf{S}_\bullet \rangle\ \tilde{d}W = \mathcal{H}_v(t, \varrho, \nabla \mathsf{S}_\bullet)\, dt \tag{28}$$

where $\tilde{d}W(t) := W(t) - W(t - dt)$ is the past-pointing Itô differential.

Proof. For simplicity, we suppress the ϱ dependences. We shall take $\varepsilon > 0$ to be infinitesimal and recast (27) in the form

$$\begin{aligned}&\left[\mathsf{S}_\bullet\left(t + \frac{1}{2}\varepsilon\right) - \mathsf{S}_\bullet\left(t - \frac{1}{2}\varepsilon\right)\right] - \mathcal{H}_v(t, \delta\mathsf{S}_\bullet(t))\,\varepsilon \\ &+ \langle \sigma, \delta\mathsf{S}_\bullet(t)\rangle \left[W\left(t + \frac{1}{2}\varepsilon\right) - W\left(t - \frac{1}{2}\varepsilon\right)\right] = o(\varepsilon).\end{aligned}$$

In time-reversed notations, this becomes

$$\begin{aligned}&\left[\tilde{\mathsf{S}}_\bullet\left(\tau - \frac{1}{2}\varepsilon\right) - \tilde{\mathsf{S}}_\bullet\left(\tau + \frac{1}{2}\varepsilon\right)\right] - \tilde{\mathcal{H}}_v\left(t, -\delta\tilde{\mathsf{S}}_\bullet(\tau)\right)\varepsilon \\ &+ \left\langle \sigma, \delta\tilde{\mathsf{S}}_\bullet(\tau)\right\rangle \left[\tilde{W}\left(\tau - \frac{1}{2}\varepsilon\right) - \tilde{W}\left(\tau + \frac{1}{2}\varepsilon\right)\right] = o(\varepsilon),\end{aligned}$$

where $\tilde{\mathcal{H}}_v(t, \varrho, X) = \mathcal{H}_v(t, \varrho, -X)$. We then have the forward-time equation

$$\begin{aligned}&\left[\tilde{\mathsf{S}}_\bullet\left(\tau + \frac{1}{2}\varepsilon\right) - \tilde{\mathsf{S}}_\bullet\left(\tau - \frac{1}{2}\varepsilon\right)\right] + \tilde{\mathcal{H}}_v\left(t, -\delta\tilde{\mathsf{S}}_\bullet(\tau)\right)\varepsilon \\ &+ \left\langle \sigma, \delta\tilde{\mathsf{S}}_\bullet(\tau)\right\rangle \left[\tilde{W}\left(\tau + \frac{1}{2}\varepsilon\right) - \tilde{W}\left(\tau - \frac{1}{2}\varepsilon\right)\right] = o(\varepsilon),\end{aligned}$$

and using the Itô-Stratonovich transformation

$$\begin{aligned}
& \left\langle \sigma, \delta\tilde{\mathsf{S}}\left(\tau\right)\right\rangle \left[\tilde{W}\left(\tau+\frac{1}{2}\varepsilon\right)-\tilde{W}\left(\tau-\frac{1}{2}\varepsilon\right)\right]+o\left(\varepsilon\right) \\
= & \left\langle \sigma, \delta\tilde{\mathsf{S}}\left(\tau\right)\right\rangle \left[\tilde{W}\left(\tau+\varepsilon\right)-\tilde{W}\left(\tau\right)\right]-\frac{1}{2}\left\langle \sigma, \delta\left\langle \sigma, \delta\tilde{\mathsf{S}}\left(\tau\right)\right\rangle\right\rangle \varepsilon,
\end{aligned}$$

we get by substitution

$$\begin{aligned}
& \left[\tilde{\mathsf{S}}\left(\tau+\frac{1}{2}\varepsilon\right)-\tilde{\mathsf{S}}\left(\tau-\frac{1}{2}\varepsilon\right)\right]-\frac{1}{2}\left\langle \sigma, \delta\left\langle \sigma, \delta\tilde{\mathsf{S}}\left(\tau\right)\right\rangle\right\rangle \varepsilon \\
& +\tilde{\mathcal{H}}_{v}\left(t,-\delta\tilde{\mathsf{S}}\left(\tau\right)\right)\varepsilon+\left\langle \sigma, \delta\tilde{\mathsf{S}}\left(\tau\right)\right\rangle \left[\tilde{W}\left(\tau+\varepsilon\right)-\tilde{W}\left(\tau\right)\right]=o\left(\varepsilon\right).
\end{aligned}$$

or, in the backward form for the original $\mathsf{S}_{\bullet}\left(t,\varrho\right)=\tilde{\mathsf{S}}_{\bullet}\left(T-t,\varrho\right)$,

$$\begin{aligned}
& \left[\mathsf{S}\left(t+\frac{1}{2}\varepsilon\right)-\mathsf{S}\left(t-\frac{1}{2}\varepsilon\right)\right]+\frac{1}{2}\left\langle \sigma, \delta\left\langle \sigma, \delta\mathsf{S}\left(t\right)\right\rangle\right\rangle \varepsilon \\
& -\mathcal{H}_{v}\left(t,\delta\mathsf{S}\left(t\right)\right)\varepsilon+\left\langle \sigma, \delta\mathsf{S}\left(t\right)\right\rangle \left[W\left(t\right)-W\left(t-\varepsilon\right)\right]=o\left(\varepsilon\right).
\end{aligned}$$

In the differential form this clearly is the same as (28). ■

If we denote by $\mathbb{E}^{(t_0,\varrho_0)}$ expectation (conditional on $\varrho_\omega\left(t_0\right)=\varrho_0$), then $\mathbb{E}^{(t_0,\varrho_0)}\left[\left\langle \sigma, \delta\mathsf{S}_{\bullet}\left(t\right)\right\rangle\ \tilde{d}W\left(t\right)\right]=0$ since $\mathsf{S}_{\bullet}\left(t\right)$, and its derivatives, are independent of the mean-zero past-point Itô differentials. We then have as a corollary that the averaged cost $\mathsf{S}\left(t,\varrho\right)$ defined by

$$\mathsf{S}\left(t_0,\varrho_0\right):=\mathbb{E}^{(t_0,\varrho_0)}\left[\mathsf{S}_{\bullet}\left(t_0,\varrho_0\right)\right]$$

will satisfy the equivalent diffusive Hamilton-Jacobi equation

$$\frac{\partial\mathsf{S}}{\partial t}+\frac{1}{2}\left\langle \sigma, \delta\left\langle \sigma, \delta\mathsf{S}\right\rangle\right\rangle=\mathcal{H}_{v}\left(t,\varrho,\delta\mathsf{S}\right) \tag{29}$$

which is the Hörmander form of the Bellman equation for optimal cost $\mathsf{S}\left(t,\varrho\right)$.

6 Linear-State Cost

A special case is applied to quantum mechanics when $\mathsf{C}\left(t,u,\varrho\right)$ and $\mathsf{S}\left(\varrho\right)$ are both linear (18) in the state ϱ with quadratic dependence of C on u. Let us specify for simplicity to a time-independent cost observable with control parameter $u=\left(u^1,\cdots,u^n\right)\in\mathbb{R}^n$ and having a quadratic dependence of the form (Einstein index notation!)

$$C\left(u\right)=\frac{1}{2}g_{\alpha\beta}u^{\alpha}u^{\beta}+u^{\alpha}F_{\alpha}+C_0$$

where $(g_{\alpha\beta})$ are the components of a symmetric positive definite metric with inverse denoted $\left(g^{\alpha\beta}\right)$ and $F_1, \cdots, F_n, C_0$ are fixed bounded operators. We take control Hamiltonian operator to be

$$H(u) = u^\alpha V_\alpha$$

where $V_1, \cdots, V_n$ are fixed bounded observables. Our aim is to find the optimal value u^* for each pair (P, Q) giving a minimum to $\langle P, C(u)\rangle + \langle w(t, u, P), Q\rangle = -\mathcal{K}_w(t, u, P, Q)$: we will have

$$\begin{aligned} 0 &= \frac{\partial}{\partial u^\alpha}\left\{\langle P, C(u)\rangle + \langle w(t, u, P), Q\rangle\right\} \\ &= g_{\alpha\beta} u^\beta + \langle P, F_\alpha\rangle + \langle i[P, V_\alpha], Q\rangle . \end{aligned}$$

Thus the optimal control $u^*(P, Q)$ is given by the components

$$u^\alpha = -g^{\alpha\beta}\left\langle P, F_\beta + \frac{1}{i}[Q, V_\beta]\right\rangle .$$

This yields a unique point of infimum and on substituting we determine that

$$\begin{aligned} \mathcal{H}_w(P, Q) &= \frac{1}{2} g^{\alpha\beta}\left\langle P, F_\alpha + \frac{1}{i}[Q, V_\alpha]\right\rangle\left\langle P, F_\beta + \frac{1}{i}[Q, V_\beta]\right\rangle \\ &\quad - \langle P, C_0 + \mathcal{L}_R(Q) + \mathcal{L}_L(Q)\rangle . \end{aligned}$$

As a result, the Hamilton-Jacobi-Bellman equation takes the form

$$\begin{gathered} \frac{\partial \mathsf{S}}{\partial t} - \frac{1}{2} g^{\alpha\beta}\left\langle \varrho, F_\alpha + \frac{1}{i}[\delta\mathsf{S}, V_\alpha]\right\rangle\left\langle \varrho, F_\beta + \frac{1}{i}[\delta\mathsf{S}, V_\beta]\right\rangle \\ + \langle \varrho, C_0 + \mathcal{L}_R(\delta\mathsf{S}) + \mathcal{L}_L(\delta\mathsf{S})\rangle + \frac{1}{2}\langle \sigma(\varrho) \otimes \sigma(\varrho), (\delta \otimes \delta)\mathsf{S}\rangle = 0. \end{gathered}$$

The terminal condition being that $\mathsf{S}(\varrho, T) = \langle \varrho, S\rangle$.

6.1 Controlled Qubit

Let us illustrate the above for the case of a qubit (two-state system). The problem we consider is similar to the one formulated in [11]. Denoting the Pauli spin vector by $\vec{\varsigma} = (\varsigma_x, \varsigma_y, \varsigma_z)$ with

$$\varsigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \varsigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \varsigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

we may represent each state by polarization vector $\vec{r} \in \mathbb{R}^3$ as

$$\varrho = \frac{1}{2}(1 + \vec{p}.\vec{\varsigma})$$

where $|\vec{p}| \leq 1$, while any observable takes the form

$$Q = q_0 + \vec{q}.\vec{\varsigma}$$

and we have the duality $\langle \varrho, Q \rangle = q_0 + \vec{q}.\vec{p}$. We shall write $\vec{p} = (x, y, z)$ and $\vec{q} = (q_x, q_y, p_z)$.

Let us suppose that we have maximal control of the Hamilton component of the dynamics, that is, we set

$$H(\vec{u}) = \frac{1}{2} \vec{u}.\vec{\varsigma}$$

with control variable $\vec{u} \in \mathbb{R}^3$. We also ignore the effect of the environment and take $\mathcal{L}_R \equiv 0$. For simplicity, we shall take the cost to have the form

$$\mathsf{C}(t, u, \varrho) = \frac{1}{2} |\vec{u}|^2$$

and we take the coupling of the system to the measurement apparatus to be determined by the operator

$$L = \frac{1}{2} \kappa \varsigma_z .$$

Explicitly we have

$$\langle w(t, u, \varrho), Q \rangle = \vec{u}.(\vec{p} \times \vec{q}) - \frac{1}{2} \kappa (x q_x + y q_y)$$

from which we see that the minimizing control is $\vec{u}^* = \vec{q} \times \vec{p}$ leading to the Hamiltonian function

$$\mathcal{H}_w(\vec{p}, \vec{q}) = -\frac{1}{2} |\vec{q} \times \vec{p}|^2 - \frac{1}{2} \kappa (x q_x + y q_y) .$$

Meanwhile, $\sigma(\varrho) \equiv \kappa (\varrho \varsigma_z + \varsigma_z \varrho) - \langle \varrho, 2\kappa\varsigma_z \rangle \varrho$ and so

$$\langle \sigma(\varrho), Q \rangle = -\kappa z x q_x - z y q_y + \kappa \left(1 - z^2\right) q_z .$$

With the customary abuse of notation, we write $\mathsf{S}(t, \varrho) \equiv \mathsf{S}(t, x, y, z)$. The Itô correction term, $\frac{1}{2} \langle \sigma(\varrho) \otimes \sigma(\varrho), \delta \otimes \delta \mathsf{S} \rangle$, in the HJB equation is then given by $\left(\text{with } \mathsf{S}_{xy} = \frac{\partial \mathsf{S}}{\partial x \partial y}, \text{ etc.}\right)$

$$\frac{\kappa^2}{2} \begin{pmatrix} -zx, & -zy, & 1 - z^2 \end{pmatrix} \begin{pmatrix} \mathsf{S}_{xx} & \mathsf{S}_{xy} & \mathsf{S}_{xz} \\ \mathsf{S}_{yx} & \mathsf{S}_{yy} & \mathsf{S}_{yz} \\ \mathsf{S}_{zx} & \mathsf{S}_{zy} & \mathsf{S}_{zz} \end{pmatrix} \begin{pmatrix} -zx \\ -zy \\ 1 - z^2 \end{pmatrix} .$$

Putting everything together, we find that the Hamilton-Jacobi-Bellman equation is

$$\begin{aligned} 0 \quad = \quad & \frac{\partial \mathsf{S}}{\partial t} - \frac{1}{2} \left| \vec{q} \times \vec{\nabla} \mathsf{S} \right|^2 - \frac{1}{2} \kappa \left(x \frac{\partial \mathsf{S}}{\partial x} + y \frac{\partial \mathsf{S}}{\partial y} \right) \\ & + \frac{\kappa^2}{2} \left(x^2 z^2 \frac{\partial^2 \mathsf{S}}{\partial x^2} + y^2 z^2 \frac{\partial^2 \mathsf{S}}{\partial y^2} + \left(1 - z^2\right)^2 \frac{\partial^2 \mathsf{S}}{\partial z^2} \right. \\ & \left. + \, x y z^2 \frac{\partial^2 \mathsf{S}}{\partial x \partial y} - x z \left(1 - z^2\right) \frac{\partial^2 \mathsf{S}}{\partial x \partial z} - y z \left(1 - z^2\right) \frac{\partial^2 \mathsf{S}}{\partial y \partial z} \right) . \end{aligned}$$

7 Discussion

In our analysis we have sought to think of the quantum state of a controlled system (that is, its von Neumann density matrix) in the same spirit as classical control engineers think about the state of the system. The advantage of this is that all the quantum features of the problem are essentially tied up in the state: once the measurements have been performed the information obtained can be treated as essentially classical, as can the problem of using this information to control the system in an optimal manner. The disadvantage is that we have to deal with a stochastic differential equation on the infinite dimensional space of quantum states. Nevertheless, the Bellman principle can then be applied in much the same spirit as for classical states and we are able to derive the corresponding Hamilton-Jacobi-Bellman theory for a wider class of cost functionals than traditionally considered in the literature. When restricted to a finite-dimensional representation of the state (on the Bloch sphere for the qubit) with the cost being a quantum expectation, we recover the class of Bellman equations encountered as standard in quantum feedback control.

Acknowledgment
We would like to thank Luc Bouten, Ramon van Handel, Hideo Mabuchi, Aubrey Truman for useful discussions. J.G. would like to acknowledge the support of EPSRC research grant GR/R78404/01, and V.P.B. acknowledges EEC support through the ATESIT project IST-2000-29681 and the RTN network QP&Applications.

References

[1] M. Armen, J. Au, J. Stockton, A. Doherty, and H. Mabuchi. Adaptive homodyne measurement of optical phase, Phys. Rev. A 89:133602, (2002)

[2] V.P. Belavkin, Theory of the control of observable quantum systems. Autom. Remote Control, 44: 178-188, (1983)

[3] V.P.Belavkin, Nondemolition measurements, nonlinear filtering and dynamical programming of quantum stochastic processes. In: Modelling and Control of Systems (Lecture Notes in Control and Information Sciences), ed A Blaquiere, 121: 381–92 (Berlin: Springer, 1988)

[4] V.P.Belavkin, A new wave equation for continuous nondemolition measurement. Phys. Lett. A, 140: 355–8 (1989).

[5] V.P.Belavkin, Stochastic posterior equations for quantum nonlinear filtering. Probability Theory and Mathematical Statistics, ed B Grigelionis, 1: 91–109 (Vilnius: VSP/Mokslas, 1990).

[6] V.P. Belavkin, Quantum stochastic calculus and quantum nonlinear filtering. Journal of Multivariate Analysis, 42: 171-201, (1992)

[7] V.P. Belavkin, Quantum continual measurements and a posteriori collapse on CCR. Commun. Math. Phys., 146, 611-635, (1992)

[8] V.P.Belavkin, Measurement, filtering and control in quantum open dynamical systems. Rep. Math. Phys. 43: 405–425 (1999).

[9] R. Bellman, Dynamic Programming, Princeton University Press (1957)

[10] J.M. Bismut, Mecanique Aléatoire. Lecture Notes in Mathematics 866, Springer-Verlag,Berlin, (1981)

[11] L. Bouten, S. Edwards, V.P. Belavkin, Bellman equations for optimal feedback control of qubit states, arXiv:quant-ph/0407192v1 (2004)

[12] L. Bouten, M. Guţă, H. Maassen, Stochastic Schrödinger equations, J. Phys. A: Math. Gen., (37): 3189-3209, (2004)

[13] M.H.A. Davis, Linear Estimation and Stochastic Control, Chapman and Hall Publishers (1977)

[14] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, Encyclopedia of Mathematics and its Applications, Cambridge University Press (1992)

[15] A. Doherty, S. Habib, K. Jacobs, H. Mabuchi, and S. Tan. Quantum feedback and classical control theory. Phys. Rev. A, 62:012105 (2000)

[16] C.W. Gardiner, Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences, Springer, Berlin (2004)

[17] C.W. Gardiner and P. Zoller, Quantum Noise. Springer, Berlin (2000)

[18] J. Geremia, J. Stockton, A. Doherty, and H. Mabuchi. Quantum Kalman filtering and the Heisenberg limit in atomic magnetometry. Phys. Rev. Lett., 91:250801 (2003)

[19] J. Gough, A. Sobolev, Stochastic Schrödinger equations as limit of discrete filtering, Open Sys. & Information Dyn., 11,235-255, (2004)

[20] J. Gough, A. Sobolev, Continuous measurement of canonical observables and limit stochastic Schrödinger equations, Phys. Rev. A 69, 032107 (2004)

[21] R. van Handel, J, Stockton, and H. Mabuchi. Feedback control of quantum state reduction. arXiv:quant-ph/0402136, (2004)

[22] M.R. James, Risk-sensitive optimal control of quantum systems, Phys. Rev. A., 69: 032108, (2004)

[23] K.R. Parthasarathy, An Introduction to Quantum Stochastic Calculus, Birkhäuser, Berlin (1992)

[24] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, E.F. Mishchenko, The Mathematical Theory of Optimal Processes, John Wiley & Sons, (1962)

[25] Scott A.J., Milburn, G.J. *Quantum nonlinear dynamics of continuously measured systems* Phys. Rev. A, Vol. 63, 042101; also arXiv:quant-ph/0008108 (2001)

[26] Smolyanov, O.G., Truman, A. *Schrödinger-Belavkin equations and associated Kolmogorov and Lindblad equations* Theor. Math, Physics, Vol. 120, No. 2, 973-984 (1993)

[27] Stratonovich, R.L.: A new representation of stochastic integrals and equations, SIAM J. Control, **4**, 362-371 (1966)

[28] A. Truman, H.Z. Zhao, The stochastic Hamilton-Jacobi equation, stochastic heat equations and Schrödinger equations, in Stochastic Analysis and Applications, D. Elworthy, I.M. Davies, A. Truman (Eds.), World Scientific Press, 441-464 (1996)

[29] M. Wiseman and G.J. Milburn, Quantum theory of optical feedback via homodyne detection. Phys. Rev. Lett. 70(5):548-551 (1993)

[30] E. Wong, M. Zakai, On the relationship between ordinary and stochastic differential equations, Int.. J. Eng. Sci., 3, pp. 213-229 (1965)